

0590  
1/16/02  
RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/899,046

OIPE

DATE: 02/13/2002  
TIME: 21:11:26

INPUT SET: S36759.raw

# 3

This Raw Listing contains the General  
Information Section and up to the first 5 pages.

## SEQUENCE LISTING

## (1) General Information:

## (i) APPLICANT:

(A) NAME: Innogenetics sa.  
(B) STREET: Industriepark Zwijnaarde 7, box 4  
(C) CITY: Ghent  
(E) COUNTRY: Belgium  
(F) POSTAL CODE (ZIP): B-9052  
(G) TELEPHONE: 00 32 9 241 07 11  
(H) TELEFAX: 00 32 9 241 07 99

**ENTERED**

(ii) TITLE OF INVENTION: New sequences of hepatitis C virus  
genotypes for diagnosis, prophylaxis and therapy.

(iii) NUMBER OF SEQUENCES: 270

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/899,046  
(B) FILING DATE:

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/362,455  
(B) FILING DATE:

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 213 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/899,046DATE: 02/13/2002  
TIME: 21:11:27

INPUT SET: S36759.raw

47 (iii) ANTI-SENSE: NO  
48  
49 (vii) IMMEDIATE SOURCE:  
50 (B) CLONE: BR34-4-20  
51  
52 (ix) FEATURE:  
53 (A) NAME/KEY: CDS  
54 (B) LOCATION: 1..213  
55  
56 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
57  
58 CTC ACG GAA CGG CTT TAC TGC GGG GGC CCT ATG TTC AAC AGC AAG GGG 48  
59 Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly  
60 1 5 10 15  
61  
62 GCC CAG TGT GGT TAT CGC CGC TGC CGT GCC AGT GGA GTT CTG CCT ACC 96  
63 Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr  
64 20 25 30  
65  
66 AGC TTC GGC AAC ACA ATC ACT TGC TAC ATC AAG GCC ACA GCG GCT GCA 144  
67 Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala  
68 35 40 45  
69  
70 AGG GCC GCA GGC CTC CGG AAC CCG GAC TTT CTT GTC TGC GGA GAT GAT 192  
71 Arg Ala Ala Gly Leu Arg Asn Pro Asp Phe Leu Val Cys Gly Asp Asp  
72 50 55 60  
73  
74 CTG GTC GTG GTG GCT GAG AGT 213  
75  
76  
77  
78 Leu Val Val Val Ala Glu Ser  
79 65 70  
80  
81 (2) INFORMATION FOR SEQ ID NO: 2:  
82  
83 (i) SEQUENCE CHARACTERISTICS:  
84 (A) LENGTH: 71 amino acids  
85 (B) TYPE: amino acid  
86 (D) TOPOLOGY: linear  
87  
88 (ii) MOLECULE TYPE: protein  
89  
90 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
91  
92 Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly  
93 1 5 10 15  
94  
95 Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr  
96 20 25 30  
97  
98 Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala  
99 35 40 45

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/899,046DATE: 02/13/2002  
TIME: 21:11:27

INPUT SET: S36759.raw

100  
101 Arg Ala Ala Gly Leu Arg Asn Pro Asp Phe Leu Val Cys Gly Asp Asp  
102 50 55 60  
103  
104 Leu Val Val Val Ala Glu Ser  
105 65 70  
106  
107 (2) INFORMATION FOR SEQ ID NO: 3:  
108  
109 (i) SEQUENCE CHARACTERISTICS:  
110 (A) LENGTH: 213 base pairs  
111 (B) TYPE: nucleic acid  
112 (C) STRANDEDNESS: single  
113 (D) TOPOLOGY: linear  
114  
115 (ii) MOLECULE TYPE: cDNA  
116  
117 (vii) IMMEDIATE SOURCE:  
118 (B) CLONE: BR36-23-18  
119  
120 (ix) FEATURE:  
121 (A) NAME/KEY: CDS  
122 (B) LOCATION: 1..213  
123  
124 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
125  
126 CTC ACG GAA CGG CTT TAC TGC GGG GGC CCT ATG TTC AAC AGC AAG GGG 48  
127 Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly  
128 1 5 10 15  
129  
130 GCC CAG TGT GGT TAT CGC CGC TGC CGT GCC AGT GGA GTT CTG CCT ACC 96  
131 Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr  
132 20 25 30  
133  
134 AGC TTC GGC AAC ACA ATC ACT TGC TAC ATC AAG GCC ACA GCG GCT GCA 144  
135 Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala  
136 35 40 45  
137  
138 AGG GCC GCA GGC CTC CGG AAC CCG GAC TTT CTT GTC TGC GGA GAT GAT 192  
139 Arg Ala Ala Gly Leu Arg Asn Pro Asp Phe Leu Val Cys Gly Asp Asp  
140 50 55 60  
141  
142 CTG GTC GTG GTG GCT GAG AGT 213  
143 Leu Val Val Val Ala Glu Ser  
144 65 70  
145  
146 (2) INFORMATION FOR SEQ ID NO: 4:  
147  
148 (i) SEQUENCE CHARACTERISTICS:  
149 (A) LENGTH: 71 amino acids  
150 (B) TYPE: amino acid  
151 (D) TOPOLOGY: linear  
152

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/899,046

DATE: 02/13/2002  
TIME: 21:11:27

INPUT SET: S36759.raw

153 (ii) MOLECULE TYPE: protein

154

155 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

156

157 Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly

158 1 5 10 15

159

160 Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr

161 20 25 30

162

163 Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala

164 35 40 45

165

166 Arg Ala Ala Gly Leu Arg Asn Pro Asp Phe Leu Val Cys Gly Asp Asp

167 50 55 60

168

169 Leu Val Val Val Ala Glu Ser

170 65 70

171

172 (2) INFORMATION FOR SEQ ID NO: 5:

173

174 (i) SEQUENCE CHARACTERISTICS:

175 (A) LENGTH: 213 base pairs

176 (B) TYPE: nucleic acid

177 (C) STRANDEDNESS: single

178 (D) TOPOLOGY: linear

179

180 (ii) MOLECULE TYPE: cDNA

181

182 (iii) HYPOTHETICAL: NO

183

184 (iii) ANTI-SENSE: NO

185

186 (vii) IMMEDIATE SOURCE:

187 (B) CLONE: BR36-23-18

188

189 (ix) FEATURE:

190 (A) NAME/KEY: CDS

191 (B) LOCATION: 1..213

192

193 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

194

195 CTC ACG GAG CGG CTT TAC TGC GGG GGC CCT ATG TTT AAC AGC AAG GGG 48

196 Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly

197 1 5 10 15

198

199 GCC CAG TGT GGT TAT CGC CGT TGC CGT GCC AGT GGA GTT CTG CCT ACC 96

200 Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr

201 20 25 30

202

203 AGC TTC GGC AAC ACA ATC ACT TGT TAC ATC AAA GCC ACA GCG GCC GCA 144

204 Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala

205 35 40 45

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/899,046DATE: 02/13/2002  
TIME: 21:11:28

INPUT SET: S36759.raw

206  
207 AAA GCC GCA GGC CTC CGG AGC CCG GAC TTT CTT GTC TGC GGA GAT GAT 192  
208 Lys Ala Ala Gly Leu Arg Ser Pro Asp Phe Leu Val Cys Gly Asp Asp  
209 50 55 60  
210

211  
212  
213 CTG GTC GTG GTG GCT GAG AGT 213  
214 Leu Val Val Val Ala Glu Ser  
215 65 70  
216  
217

218 (2) INFORMATION FOR SEQ ID NO: 6:  
219

220 (i) SEQUENCE CHARACTERISTICS:  
221 (A) LENGTH: 71 amino acids  
222 (B) TYPE: amino acid  
223 (D) TOPOLOGY: linear  
224

225 (ii) MOLECULE TYPE: protein  
226

227 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
228

229 Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly  
230 1 5 10 15  
231

232 Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr  
233 20 25 30  
234

235 Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala  
236 35 40 45  
237

238 Lys Ala Ala Gly Leu Arg Ser Pro Asp Phe Leu Val Cys Gly Asp Asp  
239 50 55 60  
240

241 Leu Val Val Val Ala Glu Ser  
242 65 70  
243

244 (2) INFORMATION FOR SEQ ID NO: 7:  
245

246 (i) SEQUENCE CHARACTERISTICS:  
247 (A) LENGTH: 213 base pairs  
248 (B) TYPE: nucleic acid  
249 (C) STRANDEDNESS: single  
250 (D) TOPOLOGY: linear  
251

252 (ii) MOLECULE TYPE: cDNA  
253

254 (iii) HYPOTHETICAL: NO  
255

256 (iii) ANTI-SENSE: NO  
257  
258

**SEQUENCE VERIFICATION REPORT**  
**PATENT APPLICATION US/09/899,046**DATE: 02/13/2002  
TIME: 21:11:28**INPUT SET: S36759.raw**

Line	Error	Original Text
5	Mandatory Value Not Present	(i) APPLICANT:
6	Unknown or Misplaced Identifier	(A) NAME: Innogenetics sa.
7	Unknown or Misplaced Identifier	(B) STREET: Industriepark Zwijnaarde 7, box 4
8	Unknown or Misplaced Identifier	(C) CITY: Ghent
9	Unknown or Misplaced Identifier	(E) COUNTRY: Belgium
10	Unknown or Misplaced Identifier	(F) POSTAL CODE (ZIP): B-9052
11	Unknown or Misplaced Identifier	(G) TELEPHONE: 00 32 9 241 07 11
12	Unknown or Misplaced Identifier	(H) TELEFAX: 00 32 9 241 07 99

PAGE: 1

**SEQUENCE MISSING ITEM REPORT**  
**PATENT APPLICATION US/09/899,046**

DATE: 02/13/2002  
TIME: 21:11:28

*INPUT SET: S36759.raw*

ADDRESSEE  
STREET  
CITY  
STATE  
COUNTRY  
ZIP  
CORRESPONDENCE ADDRESS  
CLASSIFICATION

SEQUENCE CORRECTION REPORT  
PATENT APPLICATION US/09/899,046

DATE: 02/13/2002

TIME: 21:11:28

INPUT SET: S36759.raw

Line	Original Text	Corrected Text
4031	(C) INDIVIDUAL ISOLATE: BR33 and BR36	(C) INDIVIDUAL ISOLATE: BR33 and BR36
4055	(C) INDIVIDUAL ISOLATE: HD10	(C) INDIVIDUAL ISOLATE: HD10
4080	(C) INDIVIDUAL ISOLATE: BR36	(C) INDIVIDUAL ISOLATE: BR36
4104	(C) INDIVIDUAL ISOLATE: BR36	(C) INDIVIDUAL ISOLATE: BR36
4131	(C) INDIVIDUAL ISOLATE: HD10	(C) INDIVIDUAL ISOLATE: HD10
4182	(C) INDIVIDUAL ISOLATE: BR36	(C) INDIVIDUAL ISOLATE: BR36
4209	(C) INDIVIDUAL ISOLATE: HD10	(C) INDIVIDUAL ISOLATE: HD10
5299	(C) INDIVIDUAL ISOLATE: GB358	(C) INDIVIDUAL ISOLATE: GB358
5321	(C) INDIVIDUAL ISOLATE: GB549	(C) INDIVIDUAL ISOLATE: GB549
5343	(C) INDIVIDUAL ISOLATE: GB809	(C) INDIVIDUAL ISOLATE: GB809
5364	(C) INDIVIDUAL ISOLATE: GB358	(C) INDIVIDUAL ISOLATE: GB358
5387	(C) INDIVIDUAL ISOLATE: GB549	(C) INDIVIDUAL ISOLATE: GB549
5408	(C) INDIVIDUAL ISOLATE: GB809	(C) INDIVIDUAL ISOLATE: GB809
5430	(C) INDIVIDUAL ISOLATE: GB358	(C) INDIVIDUAL ISOLATE: GB358
5453	(C) INDIVIDUAL ISOLATE: GB549	(C) INDIVIDUAL ISOLATE: GB549
5475	(C) INDIVIDUAL ISOLATE: GB809	(C) INDIVIDUAL ISOLATE: GB809
5497	(C) INDIVIDUAL ISOLATE: GB358	(C) INDIVIDUAL ISOLATE: GB358
5518	(C) INDIVIDUAL ISOLATE: GB549	(C) INDIVIDUAL ISOLATE: GB549
5539	(C) INDIVIDUAL ISOLATE: GB809	(C) INDIVIDUAL ISOLATE: GB809
5561	(C) INDIVIDUAL ISOLATE: GB358 and GB809	(C) INDIVIDUAL ISOLATE: GB358 and GB809
5584	(C) INDIVIDUAL ISOLATE: GB549	(C) INDIVIDUAL ISOLATE: GB549
5606	(C) INDIVIDUAL ISOLATE: GB549	(C) INDIVIDUAL ISOLATE: GB549